# Information Theory EITN45

STEFAN HÖST

# Personnel

Lecturer: Stefan Höst

Exercises: Umar Farooq

Secretariat: Anne Andersson

Webb: `www.eit.lth.se/course/eitn45`
`elearning.eit.lth.se`

# Claude E. Shannon

1916–2001

- *A Symbolic Analysis of Relay and Switching Circuits*, Master's thesis at MIT, 1937
- *A mathematical theory of communication*, Bell System Technical Journal, 1948

# Questions in information theory

### Two important questions in information theory

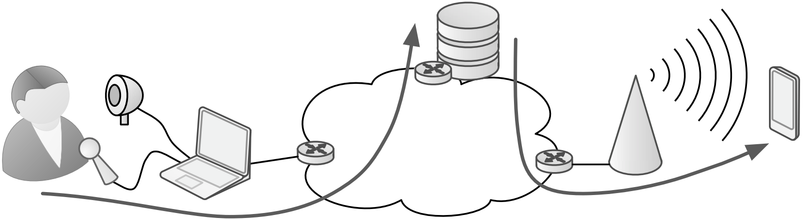
- What is information?
- What is communication?

### Shannon wrote

> *"The fundamental problem of communication is that of reproducing at one point exactly or approximately a message selected at another point."*

A “point” means “place and time”

# A communication situation



# What is information?

## A Turing Machine built using LEGO

The Turing Machine
According to Wikipedia:
"A Turing machine is a device that manipulates symbols on a strip of tape according to a table of rules. Despite its simplicity, a Turing machine can be adapted to simulate the logic of any computer algorithm, and is particularly useful in explaining the functions of a CPU inside a computer."
Unlike what the name suggests, it is not a physical machine but rather a theoretical r
mental capabilities of practically all computers in use today. This means that if somet
done on a Turing machine. This makes it a great model for scientists to use to discover
and also to show to a broad audience how a computer fundamentally operates.
However, abstract models are just that, an abstraction of something. In order to reall
computer is, we have developed a physical implementation of the Turing machine, usir
Our LEGO Turing machine uses a tape based on a classic interpretation of compute
sensor to determine the value of a switch: if the switch is on, the sensor will see the
turned off, the sensor will see the white colour of the LEGO beam, making it possib
rotating beam mounted above the tape can flip the switch in both directions.
Alan Turing's original model has an infinite tape, but LEGO had a slight problem supp
size to 32 positions. Our LEGO Turing machine only uses automatic components that are part of a single LEGO Mindstorms NXT set: one NXT brick, three electric motors and one color sensor. The final model contains parts of the NXT 2.0 set as well as a bunch of parts (mostly large beams, but also a cork screw and a set of gear racks) from two other LEGO Technic sets. We believe a slightly smaller version can be constructed using only one Mindstorms NXT 2.0 set (8547-1) and the medium-sized Technic set Mobile Crane (8053-1), which are both currently available.
[...]

## How much information?

- Text: 3363 Byte
- Zip: 1759 Byte
- Docx: 123745 Byte

# Communication theory

## Short history

**1928** Nyquist realised that a signal with bandwidth $W$ Hz and a durability of $T$ seconds can not contain more than $2WT$ distinguishable pulses.

- Certain Topics in Telegraph Transmission Theory, Bell System Technical Journal.

**1928** Hartley argued that noise limits the number of possible messages transmitted. Without noise there would not be a problem.

- Transmission of Information, Bell System Technical Journal.

**1948** Shannon wrote that communication is essential digital, and derived limits for compression rate and transmission rate.

- A Mathematical Theory of Communication, Bell System Technical Journal.

# Hartley's information measure

## Requirements for information

A symbol can provide information only if there are alternatives, i.e. only if it is the outcome from a random variable.

## Hartley's information measure

Let $X$ have $k$ possible outcomes. Then

$$I_H(X) = \log_b k$$

# Information measure for event

## Shannon's information measure

Compare conditional and unconditional probability

## Definition (Mutual Information)

The information about event $A$ from event $B$, denoted $I(A;B)$, is

$$I(A;B) = \log_b \frac{P(A|B)}{P(A)},$$

where it is assumed that $P(A) \neq 0$ and $P(B) \neq 0$.

In the continuation $b = 2$, which gives the unit bit (binary digit).
$\log_2 x = \frac{\log_b x}{\log_b 2}$ (MATLAB:log2(x))

# Information measure – Event

**Theorem**

*The mutual information is symmetric, i.e.*

$$I(A;B) = I(B;A)$$

We get the same amount of information about $A$ by observing $B$ as we get about $B$ by observing $A$. Therefore, $I(A;B)$ is called the mutual information between the events $A$ and $B$.

# Information measure – Event

**Theorem**

*The mutual information lies in the interval*

$$-\infty \leq I(A;B) \leq -\log P(A)$$

*with equality to the left if* $P(A|B) = 0$*, and with equality to the right if* $P(A|B) = 1$*.*

If $I(A;B) = 0$ the events $A$ and $B$ are statistically independent.

# Information measure – Event

**Definition (Self Information)**

The self information in the event $A$ is defined as

$$I(A) = I(A; A) = \log \frac{P(A|A)}{P(A)} = -\log P(A)$$

That is, $-\log P(A)$ is the amount of information needed to determine that the event $A$ has occurred.



LUND UNIVERSITY

# Information measure

## Definition (Entropy)

The entropy, which is a measure of the uncertainty of the random variable outcome, is

$$H(X) = E\big[I(X = x)\big] = -\sum_x p(x)\log p(x)$$

In derivations, we use the convention that $0\log 0 = 0$.

## Uncertainty can not be negative

Since $-\log p \geq 0$ for $0 \leq p \leq 1$,

$$H(X) \geq 0$$



LUND UNIVERSITY

# Example

## Example

Given a coin that could be counterfeit. The outcome from one flip has the sample space $\Omega = \{\text{Head}, \text{Tail}\}$. Denote the probabilities for the outcome by

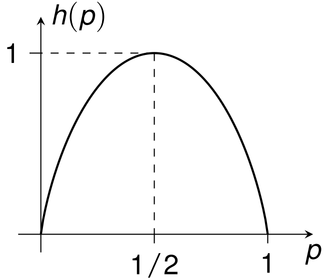
$$P(\text{Head}) = p$$
$$P(\text{Tail}) = 1 - p$$

# The binary entropy function

**Definition**

The binary entropy function is defined as

$$h(p) = -p \log p - (1-p)\log(1-p)$$



# Bounds on $H(X)$

**Theorem**

*If $X \in \{x_1, x_2, \dots, x_k\}$, then*

$$0 \leq H(X) \leq \log k$$

*with equality to the left if and only if there exists some $i$ where $p(x_i) = 1$, and with equality to the right if and only if $p(x_i) = 1/k$ for all $i = 1, 2, \dots, k$.*

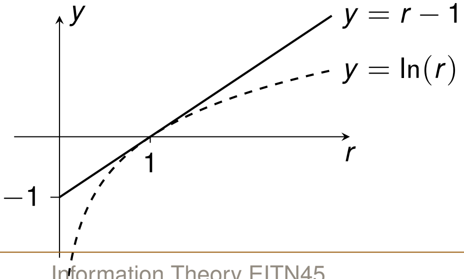
# IT inequality

**Lemma**

*For every positive real number r*

$$\log(r) \leq (r-1)\log(e)$$

*with equality if and only if* $r = 1$.






LUND UNIVERSITY